

Accelerated Strategic Computing Initiative (ASCI) : Driving the Need for the Terascale Simulation Facility (TSF)



Anna Maria Bailey, PE
TSF Design and Construction Manager
Electrical Engineer
Energy 2002 Workshop and Exposition
Palm Springs, CA - June 2- 5



Lawrence Livermore National Laboratory, P. O. Box 808, Livermore, CA 94551

Stockpile Stewardship Program

ASCI Mission - Satisfy the Computing Requirements for Stockpile Stewardship



Goal

Create the computing
environment needed to carry out
mission critical simulations and
understand their implications

Objectives

- Maintain a credible nuclear deterrent with a "zero-yield" nuclear test ban
- Support the Comprehensive Test Ban policy
- Ensure the effectiveness of science-based stockpile stewardship
- Perform high-fidelity physics simulation drivers



- 10⁹ cells
- 3D

- Turbulence
- Mix
- Transport
- Estimated 10¹⁴⁻¹⁶ Flops near-term computing capacity

WBS-1.1.D.2.7

2

Four Key Points

- Modeling and simulation has an absolute critical role in the Stockpile Stewardship Program
 - The role of simulation has increased greatly since the end of nuclear testing
 - The rate of change is driven by programmatic requirements
- Meeting the needs of the Stockpile Stewardship Program requires a next-generation simulation capability
 - A simulation capability is more than just weapons codes
 - It involves improving scientific understanding in many areas
- The Stockpile Stewardship Program need for this next-generation simulation capability drives our computing requirements
 - TeraOPS computers
 - Terascale computing environment
- The Terascale Simulation Facility (TSF) is vital to LLNL's ability to field the nation's premier modeling and simulation environment for the Department of Energy

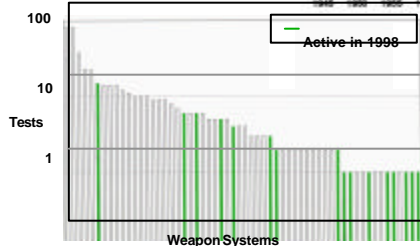
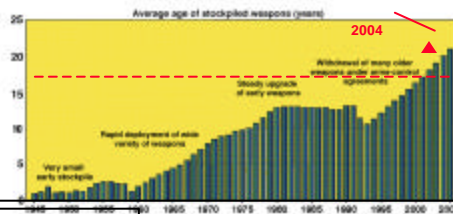
WBS-1.1.D.2.7

The FY2004 Milestone

FY2004 is an Important Deadline

By FY2004

- The average age of US nuclear weapon systems will exceed their design lifetimes



By FY2004

- Almost all of DOE's design/ test-experienced scientists will have retired.
- Those remaining will have experienced very few tests.

WBS-1.1.D.2.7

Complex modeling and simulation demands unprecedented terascale computing

- Tens of thousands of processors in hundreds of cabinets, capable of peak performance exceeding 60 TeraOPS by late 2004
- Very large aggregate memories, up to 50 terabytes (trillions of bytes)
- Archives capable of handling petabytes (quadrillions of bytes) of data
- These requirements push the boundaries of current technology
- Facilities requirements far exceed those of conventional computer centers
 - High levels of power and cooling
 - Unencumbered floor space
 - Communications infrastructure

WBS-11027

5

Simulation & Computer Science programs tie together hardware, simulations and scientists

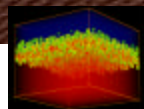
Platforms



Storage & Servers



Scientists

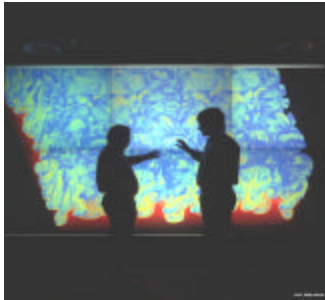


These programs provide the enabling computer and simulation technologies needed to design and implement Stockpile Stewardship simulations

WBS-11027

6

The only way to comprehend huge amounts of data is to view it



- ASCI simulations will generate results at rates of terabytes/sec
 - Library of Congress (books) » 20 TB
- Must be able to interact with (visualize, store, retrieve) multi-terabyte files

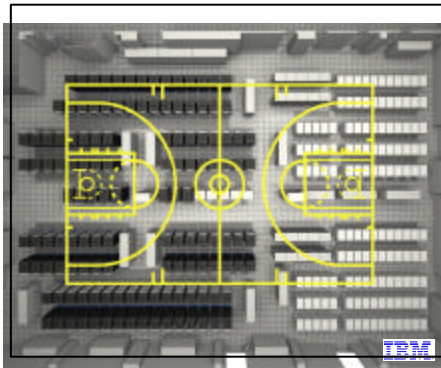
- *Visualization is the only practical approach to understanding simulation of complex nonlinear phenomena*
- *Visual channel is dominant (>50% of our neurons)*
- *TSF's Advanced Simulation Laboratory (ASL) will allow development of new visualization techniques*

WBS-1.1.D.2.7

7

Platform Development

The ASCI White system at LLNL is currently the world's fastest computer



Occupying 12,000 sq. ft., this system is only half the projected size of the first 100 TeraOPS system

WBS-1.1.D.2.7

8

The ASCI Platforms are the world's most powerful



ASCI Red

- SNL
- Intel Prime Contractor
- 3.15 TeraOPS Peak
- 2 Processors per scalable unit
- UNIX Operating System
- 9360 333 MHz Processors
- 1.2 Terabyte Memory
- 12.5 Terabyte Storage



ASCI Blue-Mountain

- LANL
- SGI Prime Contractor
- 3.072 TeraOPS Peak
- 128 Processors per scalable unit
- IRIX Operating System
- 6144 250 MHz Processors
- 1.5 Terabyte Memory
- 76 Terabyte Storage



ASCI Blue-Pacific

- LLNL
- IBM Prime Contractor
- 3.89 TeraOPS Peak
- 4 Processors per scalable unit
- AIX Operating System
- 5856 332 MHz Processors
- 2.6 Terabyte Memory
- 52.5 Terabyte Storage



ASCI White

- LLNL
- IBM Prime Contractor
- 12.2 TeraOPS Peak
- 16 Processors per scalable unit
- AIX Operating System
- 8192 310 MHz Processors
- 6 Terabyte Memory
- 150 Terabyte Storage

WBS-11027

9

ASCI High Power Density

- Estimated terascale computer power requirements have high levels of computational and mechanical power. For example, the 12 TeraOps ASCI White system installed in B451 is designed to accommodate the following:

— Power (MW) CPU + Disk	=	3.6 MW
— Mechanical (MW)	=	2.3 MW
— Total Power for System	=	5.9 MW
— Machine Room (sq. ft.)	=	12,000



Approximate power density of 350 Watts/sf.

WBS-11027

10

ASCI B-451 12 TF Space Allocation and Adjacencies are Primary Issues

- Restructure B-451 to accept unconventional terascale computing systems. The building was designed 20 years ago to accommodate smaller conventional computer systems such as CRAY platforms.
- Support Equipment for power and cooling needed to be sited adjacent to the computer room. These systems include:
 - Fan units for air cooling the computer (computer room air conditioning units)
 - Mechanical systems to provide chilled water
 - Mechanical piping between the chilled water system and the CRACs
 - Electrical power systems for the computer system
 - Electrical power systems for the mechanical systems
 - Fire protection and alarm systems
- Space requirements, power requirements, and coordination of the support systems are significant cost elements in the installation.
- Expansion of all systems for future requirements must also be part of the initial planning.
- Additional space for computer equipment provides flexibility in siting systems.

WBS-11027

11

ASCI B-451 12 TF Power Demands Drove Building Design

- Architectural requirements
 - Raised floor height increased from 18" to 30"
 - Floor space increased from 12,000 square feet to 20,000 square feet (additional space used for staging of future terascale systems)
 - Additional rooms were remodeled for electrical and boiler rooms
- Mechanical Design
 - Chilled water plant increased from 300 tons to 2025 tons
 - New chilled water piping loop installed in the computer room
 - New heating boiler plant replaced heat recovery bundle of chillers
 - 32 computer room air-conditioning units were installed
- Electrical Design
 - Three new substations added to existing electrical system in the electrical yard
 - (1) 2.5 MVA, 13.8 KV - 4.16 KV, and (2) 2000 KVA, 13.8 KV - 480/277 V
 - Major interior electrical distribution was required
 - (8) 480 V/208-120 V, 500 KVA transformers, (8) 208/120 V switchboards, and (36) 208-120 V, 400A electrical panels were installed to support the computer.

WBS-11027

12

ASCI B-451 12 TF Approaching the Problem

- Computational Fluid Dynamics (CFD) analysis was done to characterize the airflows, identify unacceptable temperature levels, and verify component arrangement.
- Building components carefully selected to provide needed airflow. (floor height, tiles, ceiling height, ceiling space)
- Support equipment carefully selected for size (to fit the existing space), arrangement, and delivery (installation prior to computer delivery).
- Underfloor and ceiling spaces were carefully analyzed for interferences between mechanical and electrical systems and to maximize space available for computer cabling and for air flow.

WBS-11027

13

ASCI B-451 12 TF Constructability Issues: Underfloor

- Congested 30" underfloor with various utilities tightly spaced around the room.
 - SRG Grounding: Required 2"
 - Leak Detection: Required 2"
 - CO2 Piping: Required 6"
 - Fire Protection: Required 4"
 - Chilled Water: Required 18"
 - Power Distribution: Required 6"



WBS-11027

14

ASCI B-451 12 TF Constructability Issues: Ceiling Space



- Congested 60" ceiling space with various utilities tightly spaced. This required a great deal of planning and maneuvering of utilities as they crossed one another in the space.
- Structural steel: Required 21"
—(Intermediate steel required 21")
- Ductwork: Required 24"
—(Supply/return cross required 48")
- Fire sprinkler: Required 4"
- Conduit: Required 6"
—(With 2" for support structures)
- Light Fixtures: Required 5"
- Ceiling Tiles: Required 4"
—(Access and removal of tiles)

WBS-1.1.D.2.7

15

The computer layout creates heat transfer problems

- Dense computer rack spacing creates narrow heat canyons. The CFD analysis was required to be performed at 2' -6", 5' -0" and 7' -6" levels to ensure that proper coverage of cooling could be achieved.
- The arrangement left limited space for heat exchange.
- The layout provides limited floor space for air distribution.



WBS-1.1.D.2.7

16

ASCI B-451 12 TF: The computer floor has limited tile space available for airflow



- The computer racks and the CRACs occupy approximately 30% of the floor space.
- About 6 square feet of floor space is available to deliver cooling capacity for each rack due to rack spacing. Cooling requirements range from 3.2 kw to 18.8 kw per rack.
- Standard perforated tiles used in most computer applications were supplemented with a large number of grated tiles to meet the projected airflow.
- Floor tile ratings:
 - Perforated tiles rated at 440 CFM with 24% free area
 - Grated tiles rated at 1970 CFM with 64% free area

WBS-11027

17

The computer room has limited wall space for electrical panels



- CRACs are typically installed at the perimeter of the computer room.
- CRACs and electrical panels compete for space along the wall.
- Installation of additional wall space through intermediate column spaces is not compatible with the computer's cabling requirements.
- Electrical wall panels were utilized in lieu of Power Distribution Units (PDUs) which are not feasible in this design due to limited space allocations. For example if a 225kVA PDU was used, approx. 150 sq. ft. is required including physical dimensions and NEC clearances. With a total of (16) PDU(s) needed for the total project load, 2400 sq.ft. or another 20% of the floor space would be unavailable.

WBS-11027

18

Electrical Design Parameters

- Provide high quality power to facility to support 5.9 MW of computational and mechanical load.
- Since UPS is too costly and space consuming, alternatives were implemented to provide redundant sources at all voltage levels to ensure that power is highly reliable.
- The design was to include flexibility for power/load swing in the distribution system for the load. The facility has the ability to dynamically redistribute power and cooling up to 5.9 MW of load.
- Power separation for improved power quality was incorporated. The mechanical loads are separated from the computational loads.
- A major component was to ensure that the design was fully functional, scalable, and flexible.

WBS-1.1.D.2.7

19

Electrical Design Parameters



- The existing facility capacity was 3MW with a 1.5MW load.
- The facility was upgraded with a 13.8kv, 7.5MW feeder to accommodate the existing and the new 5.9MW projected load.
- The electrical yard was expanded to accommodate the new substations and interconnected via overhead busduct.
- The new feeder is redundant at the 13.8kV level providing immediate backup.
- Existing and new substations are tied together with a backup to load shift in the event that one of the substations is down.
- The peak demand at LLNL today is approx. 60 MW for 1 square mile, this facility is adding 10% of the peak value in approximately 12,000 square feet.

WBS-1.1.D.2.7

20

Mechanical Design Parameters



- The chilled water system for supplying the CRACs was sized to support the computer room, the addition to the computer room, and the office building.
- Three 675-ton chillers, one 270-ton chiller, and three chilled water pumps will cycle on and off to meet the demands of computers and building loads from no load to 4.5 MW of demand.
- Each 675-ton chiller cools about 2 MW. The chiller overall size is 17 feet long, 8 feet wide and 10 feet in height.
- Chiller system is designed for a N+1 redundancy to ensure that the cooling system is properly backed up.
- Variable speed pumping and primary-secondary piping loops can be used to provide flexibility in chiller plant operation.

WBS-1.1.D.2.7

21

Mechanical Design Parameters



- 14-inch and 16-inch chilled water and condenser water piping were installed in the mechanical room between the chillers and pumps.
- Clearance under the piperack provides future capability of moving out the chillers.
- The piperack supports are 12-foot, 2 ft diameter piers with 18-inch deep wide flange structural steel for the support structure.
- Existing underground utilities were relocated to accommodate the massing of structural support.

WBS-1.1.D.2.7

22

Direct Digital Control (DDC) System is Implemented



- A DDC system is implemented to control all components of the mechanical system to ensure that the system is capable of providing the appropriate airflow for the load.

WBS-11027

23

ASCI B-451 12 TF Constructability Issues: Computer Installation



- When the computer arrived, miles of connections were accomplished with all of the utility congestion under the floor.
- The cable length was minimized so that maximum cable distance between any two nodes is not exceeded.
- Approximately 14" of cable depth is in place over the utilities beneath the raised floor with the computer completely installed.

WBS-11027

24

Flexibility in siting terascale computer equipment



- The computer room was also expanded from 12,000 sq ft to nearly 20,000 sq ft to provide flexibility in siting computer systems.
- The addition provides space for installation of very small additional computer systems concurrently with the 12-TeraOps machine and for permitting reconfiguration of the 12-TeraOps during its use.
- In the future, the expanded area will also be used to house the growing network environment.
- This building is at its computational capacity with a computer in the 20 to 30 TeraOps range. This drives the need for a new facility to house the terascale 100 TeraOps future computer.

WBS-11027

26

The Future of ASCI

The Advanced Simulation and Computing Program (ASCI) Terascale Simulation Facility (TSF) at LLNL is being constructed to provide power and space to accommodate two simultaneous 100-teraOPs-class computer systems' operations, including assessment areas and networking control areas necessary for direction and assimilation of data.

The first Laboratory building dedicated to computing to be constructed in 20 years, TSF will become the home for next-generation supercomputers, starting with delivery of the 60+ teraOPs (60 trillion operations per second) ASCI Option Purple in June 2004.

Simulation is a cornerstone of the National Nuclear Security Administration's (NNSA) Stockpile Stewardship Program to ensure the safety and reliability of the nation's nuclear weapons stockpile without testing.

WBS-11027

26

Aerial View of Future TSF Site

This preconstruction aerial view of the site for TSF shows the Drainage Retention Basin (top), Building 551W (right), and Building 451 and Trailer 4525 (center bottom).



WBS-11027

27

Terascale Simulation Facility (TSF)

The TSF will encompass 253,000 square feet. Two 128-foot by 192-foot two-story computer rooms will provide a total of 48,000 square feet of space in two rooms for computer systems. A four-story office complex will provide space for 288 scientists, engineers, and support staff.



WBS-11027

28

TSF Perspective

The TSF as seen from the parking lot located south of the building as envisioned after completion with landscaping and cars.



WBS-11027

29

TSF Site Plan

TSF will occupy the area immediately south of the Drainage Retention Basin, west of Building 551, and immediately northeast of Building 451.



WBS-11027

30

ASCI Campus



The TSF as seen from Building 451. The TSF and Building 451 will form a modern, centralized campus for high-performance computing infrastructure at LLNL. Building 451 is home to the world's fastest computer, ASCI White. The TSF will house ASCI Purple.

WBS-11027

31

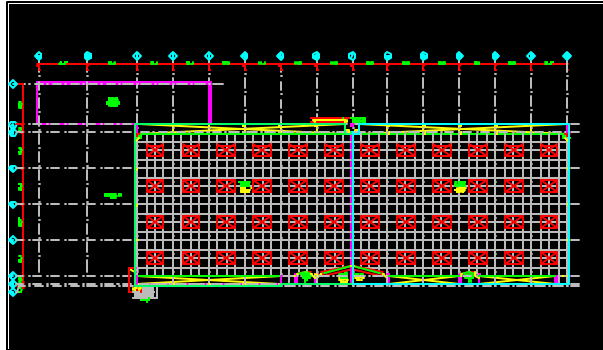
TSF Highlights

- (2) 24,000 square foot computer rooms are required to handle a 100 TeraOps computer in each room. A new system is staged in every 18 months with a 6 month overlap where both rooms are computing.
- The computer rooms must be free of obstructions such as columns, CRAC units and PDU(s) which impede on vital computational floor space. The 20,000 square foot computer room in B-451 is limited to a 20 to 30 TeraOps due to all of the above and below floor impediments.
- The raised floor must be 48" to accommodate the miles of cable connections. The 30" floor at B-451 is a limiting factor.
- The mechanical chilled water plant must be designed for a N+1 redundancy to ensure that the cooling system is properly backed up.
- The electrical system must be highly reliable and scalable to handle the load.

WBS-11027

32

TSF Computational Space



The west computer room is designed for a computational load of 6 MW with a future expansion capability to 7.5 MW. The east computer room is designed for a computational load of 3 MW with a future expansion capability to 7.5 MW.

WBS-11027

33

TSF Mechanical Approach



In order to provide adequate airflow and not impede on the computer floor space, air is produced in the mechanical room below the computer floor and distributed through the openings between the floors. This design eliminates the need for CRAC units on the floor which require an additional 8,000 square feet of computer floor space.

WBS-11027

34

TSF Electrical Approach

In order to provide adequate electrical distribution for the facility, a new 13.8kV switching station referred to as LGS-TSF must be installed to accommodate a future TSF load of 18MW. In addition to the new LGS, the following equipment must be installed in a very compact electrical yard:

(6) sectionalizing switches and associated vaults, (1) 13.8kV - 4.16kV double ended chiller substation, (15) 13.8kV - 480/120V 1000kVA pad mount transformers and associated switchboards, (1) 150kW standby generator. The (2) building load substations are to be installed inside of the building in an electrical room due to space constraints in the electrical yard.

WBS-11027

36

ASCI Projected Power Utilization Forecast

Date	ASCI Computational Electrical Level for TSF (MW)	ASCI Mechanical Electrical Level for TSF (MW)		TSF Office Tower MW	B-451 ASCI Load MW	Total
June-04	5.0	3.5		0.0	2.0	10.5
January-05	5.0	3.5		2.0	1.2	11.7
January-06	4.0	2.8		2.0	1.2	10.0
June-07	9.0	6.3		2.0	1.2	18.5
June-08	6.0	4.2		2.0	1.2	13.4

WBS-11027

37

TSF Computational Utilization Probability

Computational Electrical Level		
Date	at TSF (MW)	Probability(*)
June-04	5.0	75.0%
January-05	5.0	95.0%
January-06	4.0	95.0%
June-07	9.0	40.0%
June-08	6.0	40.0%

(*) Probability is based on the advancement and availability in technology of the processor and microprocessor design.